F.A.Q. Probabilités et Statistiques

Le présent document est un support pédagogique visant à développer l'intuition du probabiliste et statisticien en herbe. Il est en aucun cas un cours à proprement parler. Pour ne pas alourdir les réponses, nous avons fait le choix de ne pas donner les démonstration des résultats énoncés.

Question 1. *Change-t-on le problème si on met un ordre sur les variables aléatoires ? Par exemple, sur deux dés, distinguer le dé 1 du dé 2 ?*

Oui et Non ! On ajoute une hypothèse (qui change le problème) qu'il faut enlever à la fin du raisonnement (pour rétablir le problème initial). Cela peut rendre les calculs plus simples, ou plus compliqués...
<u>Ex :</u> Si on s'intéresse à compter le nombre de combinaisons de deux dés tels que les deux chiffres soient consécutifs, on peut tous les compter (dans un tableau par exemple). Cela fait 10. Ou on peut choisir de donner un ordre aux dés. Dans ce cas, on compte les combinaisons 1-2, 2-3, 3-4, 4-5, 5-6. Il faut ensuite enlever cette hypothèse d'ordre. Ici, on remarque qu'à chaque configuration on peut associer une configuration "soeur" en intervertissant les deux chiffres. On a donc bien $2 \times 5 = 10$ combinaisons.

Question 2. *Pourquoi lorsque l'on regarde un $\Omega = \{1, 2, 3, 4\}$ par exemple, n'a-t-on pas toujours $P(1) = P(2) = P(3) = P(4) = \frac{1}{4}$?*

Lorsque l'on ne précise rien, il se peut que les évènements élémentaires ne soient pas tous équiprobables.
<u>Ex :</u> Imaginons que le dé à quatre faces que l'on regarde soit pipé (truqué). Ou bien que l'on ai effacé le chiffre quatre pour le remplacer par un deux. Nous aurions créé (dans la vraie vie ! ! !) un objet qui répond à la loi de probabilité $P(1) = P(3) = \frac{1}{4}$, $P(2) = \frac{1}{2}$ et $P(4) = 0$.

Question 3. *Pourquoi compte-t-on le nombre de possibilités d'un événement au lieu de calculer des produits de probabilités ?*

Dans beaucoup de cas, (dés, tirages de boules,...) toutes les configurations élémentaires sont équiprobables. C'est-à-dire : $\forall\ \omega, \omega' \in \Omega,\ P(\omega) = P(\omega') = p$
Ce qui donne : $\sum_{\omega\in\Omega} P(\omega) = 1 \Rightarrow \#\Omega \times p = 1$ et donc $p = \frac{1}{\#\Omega}$
Donc lorsque l'on identifie chaque élément ω correspondant à un événement (et surtout que l'on sait les compter), on a

$$P(A) = \sum_{\omega\in A} P(\omega) = \sum_{\omega\in A} p = \frac{\#A}{\#\Omega}$$

Attention, ce n'est bien sûr pas le cas si tous les ω ne sont pas équiprobables ! ! !

Question 4. *Si* $\boxed{a}$ $\Omega = \{1, 2, 3, 4\}$ $\boxed{b}$ $\Omega = \{(1, 1), (1, 5)\}$ $\boxed{c}$ $\Omega = \{chaise, table\} \times \{3, 6\}$ *A quoi ressemblent les ω ? A quoi ressemble un événement ?*

$\boxed{\text{a}}$ 1 est un élément tout comme 2 par exemple. $\{2, 4\}$ est un événement, $\{1\}$ aussi.
$\boxed{\text{b}}$ $(1, 1)$ et $(1, 5)$ sont deux éléments.
$\boxed{\text{c}}$ $(chaise, 6)$ est un élément ; $(chaise, 3)$ aussi. $\{chaise, table\} \times \{6\}$ est un événement de Ω.
A chaque fois, ω appartient à Ω et un événement A est inclus dans Ω.

Question 5. *Quand ajoute-t-on deux probabilités ?*

Dans un seul cas : lorsque l'on regarde l'union de deux événements disjoints (on dit aussi incompatibles). C'est-à-dire $A \cap B = \emptyset$. Alors $P(A \cup B) = P(A \text{ ou } B) = P(A) + P(B)$

<u>Ex :</u> P("au plus un pair") $= P$("aucun pair" ou "un pair") $= P$("aucun pair" $\cup$ "un pair"). Mais comme on ne peut avoir simultanément "aucun pair" et "un pair" (les événements sont disjoints), nous avons bien :

$$P(\text{“aucun pair” ou “un pair”}) = P(\text{“aucun pair”}) + P(\text{“un pair”})$$

Dans le cas général, on peut écrire $\boxed{P(A \cup B) = P(A) + P(B) - P(A \cap B)}$
Attention, ne pas confondre incompatibles et indépendants !

Question 6. *Quand multiplie-t-on deux probabilités ?*

On se pose la question : Est-ce que la réalisation de A a une quelconque influence sur la réalisation de B ? Si la réponse est non alors les deux événements sont indépendants ! Si deux événements sont **indépendants** : nous avons $\boxed{P(A \cap B) = P(A)P(B)}$

Ex : Prenons deux dés. L'un à quatre faces comportant les chiffres 1,1,2,3. Le second à six faces comportant les chiffres 1,1,1,2,2,3. On choisit au hasard l'un des deux dés puis on le lance.
On a alors :

$$\begin{aligned} P(\text{“avoir un 1” et “choisir le dé à 6 faces”}) &= P(\text{“avoir un 1”} \cap \text{“choisir le dé à 6 faces”}) \\ &= P(\text{“avoir un 1”})P(\text{“choisir le dé à 6 faces”}) \end{aligned}$$

En effet, on se fiche de savoir si on prend le dé à six faces ou à quatre faces pour connaitre la chance d'avoir un 1. Elle de $\frac{1}{2}$ sur chaque dé ! !

Question 7. *Qu'est-ce qu'une probabilité conditionnelle ?*

Elle est définie ainsi : $P(\bullet|B) = \frac{P(\bullet \cap B)}{P(B)}$. Qu'elle opération a-t-on fait ? On a modifié la probabilité P d'origine en une autre probabilité. C'est donc une probabilité ! ! Toutes les formules sur les probabilités s'appliquent. Cela donne la probabilité d'un événement en sachant B de façon sûre. Pratique lorsque l'on veut scinder le problème.

Ex : On jete deux dés. Si on appelle X_1 le premier dé, X_2 le second et $Y = \max(X_1, X_2)$ le maximum des deux dés. Alors :

$$\begin{aligned} P(Y = 3 \text{ et } X_1 = 3) &= P(Y = 3|X_1 = 3)P(X_1 = 3) \\ &= P(X_2 = 1{,}2 \text{ ou } 3)P(X_1 = 3) \\ &= \frac{1}{2} \times \frac{1}{6} = \frac{1}{12} \end{aligned}$$

On a transformé $P(Y = 3|X_1 = 3)$ en $P(X_2 = 1{,}2 \text{ ou } 3)$ car sachant $X_1 = 3$ la seule manière d'avoir $Y = 3$ est que le deuxième dé soit plus petit.

Question 8. *Quand reconnaître une loi binomiale ?*

Uniquement dans le cas suivant : La population totale est **fixée** (n individus). On s'intéresse à la variable aléatoire X égale au **nombre** d'individus portant une certaine **propriété** (on note p la probabilité pour qu'un individu porte cette propriété). Cette variable aléatoire est alors modélisée par une loi Binomiale B(n,p).

Ex : Le nombre de garçon dans une population donnée. Le nombre d'ampoules défectueuses dans un échantillon donné. Le nombre de dé valant 6 sur un nombre de lancés donné.

Question 9. *Quand reconnaître une loi de Poisson ?*

La loi de Poisson est une loi de probabilité discrète qui décrit le comportement du **nombre** d'évènements se produisant dans un laps de **temps fixé**, si ces évènements se produisent avec une **fréquence moyenne connue** et **indépendamment du temps** écoulé depuis l'évènement précédent. La loi de Poisson est également pertinente pour décrire le nombre d'évènements dans d'autres types d'intervalles, spatiaux plutôt que temporels, comme des segments, surfaces ou volumes.

Ex : Le nombre d'appels arrivant à un standard téléphonique pendant une période donnée. Le nombre de personnes à la file d'attente d'un magasin arrivant en une durée donnée. Le nombre de poissons (!) pêchés dans un filet de pêche donné.

Question 10. *Quand reconnaître une loi géométrique ?*

Elle modélise la situation : **Tant que je perds** (avec proba $1 - p$), **je joue et dès que je gagne** (avec proba p), **je m'arrête**. La variable aléatoire X qui suit une loi géométrique est alors le nombre de fois que j'ai joué (en comptant la fois où j'ai gagné).

Ex : Le nombre de lancés de canne à pêche avant d'attraper un poisson. Le nombre de lancés de dé avant d'avoir un 6. A la pétanque, le nombre de boules à lancer avant de “prendre”. (vraiment ?)

Question 11. *Quand reconnaître une loi multinomiale ?*

C'est une généralisation de la loi binomiale. Là ou précédemment, on s'intéressait a un problème binaire (être ou ne pas être !), ici on regarde une collection d'individus qui peuvent prendre plus de 2 valeurs mais toujours en nombre fini. Chaque individu **peut prendre** i **valeurs** (de proba p_i). On regarde une population **fixée** (n individus). Alors la variable aléatoire $(X_1, \ldots, X_i)$ suit une loi multinomiale où X_i est le **nombre** d'individus dans l'état i.

Question 12. *Quand reconnaître une loi hypergéométrique ?*

Le cadre ressemble fortement à la loi binomiale ! Mais il ne faut pas les confondre. On tire aléatoirement un **nombre fixe** n de boules dans une **population déterminée** (donc non-aléatoire) comportant des boules gagnantes (au nombre de G) et des boules perdantes (au nombre de P). Alors X, le **nombre de boules gagnantes** obtenu, suit une loi hypergéométrique.
Ici les individus ne sont pas aléatoires, mais un sous-ensemble d'individus est choisi aléatoirement dans l'ensemble (loi hypergéométrique) ce qui diffère de considérer des individus aléatoires qui forment une collection fixe (loi binomiale).

Question 13. *Quelle est la différence entre X et $X(\omega)$?*

Une variable aléatoire est une fonction : $X : \Omega \to \mathbb{R}$. On connait déjà la différence entre f et $f(x)$. f est un objet mathématique formel et peut avoir des propriétés comme "être croissante", "être continue", etc... $f(x)$ est (par exemple) un réel, un réel positif, la valeur d'une intégrale, etc... Il en va de même pour X et $X(\omega)$. X peut être de loi uniforme, à support borné, symétrique, etc... $X(\omega)$ n'est quant à lui qu'un réel. C'est la méconnaissance de la dépendance en ω de $X(\omega)$ qui fait que l'on écrit très souvent X pour $X(\omega)$ mais c'est bien un abus (totalement admis) de notation.
<u>Ex :</u> $X(\omega) = \sin(\omega)$ est une variable aléatoire de loi uniforme sur $[-1, 1]$ (qui a peu d'intérêt). Les cas intéressants sont justement ceux où l'on ne connait pas la dépendance en ω.
On devrait normalement écrire $\mathbb{P}(-1 \leq X(\omega) \leq 2,7)$ au lieu de $\mathbb{P}(-1 \leq X \leq 2,7)$

Question 14. *Comment faire un plan de réponse à un exercice de statistiques ?*

On peut le construire en 4 parties :

1. Identifier le cadre
   -Soit $X_1, \ldots, X_n$ des variables **indépendantes** et de **même loi**.
   -Identifier les paramètres connus et les paramètres inconnus. (Certaines fois, la forme de la loi elle-même est inconnue.)
2. Traduire l'énoncé en quantité(s) dépendant(es) des X_i
   -Souvent $\bar{X} = \sum_{i=1}^{n} X_i$ (mais pas seulement).
   -Il y a souvent autant de quantités que de paramètres inconnus !
3. Se ramener à une loi connue (C'est la partie délicate)
   -On utilise les outils probabilistes : LGN, Inégalité de Markov, B-T, **TCL** (A connaître absolument), Approximation de loi par une autre.
   -C'est l'étape où l'on fait des calculs.
4. Conclure
   -Penser à répondre à la question posée.
   -Commenter le résultat autant que possible.

Question 15. *Qu'est ce qu'un test ? A quoi servent-ils ?*

Un test T est une fonction de la famille $X_1, \ldots, X_n$ qui renvoie 0 ou 1. Dans certains problèmes de statistique, il est plus important d'écarter une hypothèse sur le(s) paramètre(s) inconnu(s) que de chercher à avoir une bonne estimation de ce(s) paramètre(s). C'est pour cela que l'on formule des (deux) hypothèses sur la loi des X_i. Le test T vise à faire la différence entre ces deux possibilités.

Question 16. *Qu'est-ce qu'une hypothèse de test statistique ?*

Il y a deux grandes familles d'hypothèses :

1. C'est un ensemble où vit un (ou plusieurs) paramètre(s) de la loi des $X_1, \ldots, X_n$. Attention, en aucuns cas ce ne sont des évènements ! (Elles n'ont donc aucune probabilité.) La plupart du temps on note Θ l'ensemble de toutes les valeurs possibles du paramètre et θ le paramètre inconnu. Dans ce formalisme, H_0 et H_1 sont des sous-ensembles de Θ.
   Ex : Si les X_i sont des Bernoulli de paramètre p inconnu alors $\{p : p \leq \frac{1}{2}\}$ et $\{\frac{1}{4}\} \cup \{\frac{3}{4}\}$ sont deux hypothèses. En fait, tout sous-ensemble de $\Theta = [0, 1]$ convient.
2. C'est une loi particulière.
   Ex : On peut tester une loi exponentielle contre une loi normale, une loi uniforme contre une loi non-uniforme, etc... Le test du χ^2 est le test phare dans ce cas.

Question 17. *Quelle est la différence entre H_0 et H_1 ?*

On ne raisonne que dans le cas 1. précédent mais ceci s'étend évidemment au cas 2.
On se focalise sur H_0 c'est-à-dire que l'on va construire le test T tel que $\mathbb{P}_\theta(T = 1)$ est suffisamment petit (en général 5%) pour tout $\theta \in H_0$. Attention :

$$\begin{aligned} T = 1 \quad &\Longrightarrow \quad H_0 \text{ est fausse (avec forte probabilité)} \\ T = 0 \quad &\Longrightarrow \quad \text{On ne peut rien dire} \end{aligned}$$

On "espère" ensuite que l'erreur sur H_1 : $\sup_{\theta \in H_1} \mathbb{P}_\theta(T = 0)$ soit petite.
L'erreur sur H_0 : $\sup_{\theta \in H_0} \mathbb{P}_\theta(T = 1)$ est appelée **erreur de première espèce** et l'erreur sur H_1 : $\sup_{\theta \in H_1} \mathbb{P}_\theta(T = 0)$ est appelée **erreur de seconde espèce**. (Y réfléchir ! Ce sont les deux seules sources d'erreur possible.)

Question 18. *Qu'est-ce qu'une statistique de test ?*

C'est le cœur de la machine ! Un test "regarde" la valeur d'une statistique de test S et répond 0 ou 1 en fonction de celle-ci. Construire un test revient à trouver une **zone de rejet** de S.
Ex : On cherche, par exemple, à déterminer si les X_i sont d'espérance 0 (hypothèse H_0) ou d'espérance 4 (hypothèse H_1). La statistique de test serait $S = \bar{X}$. Si elle est grande, on aura tendance à répondre 1 et si elle est petite on aura tendance à répondre 0. Ici, notre test aura donc la forme $T = \mathbb{1}_{\bar{X} > t}$.
Ex : On voudrait savoir si les X_i sont de variance forte (hypothèse H_0) ou de variance faible (hypothèse H_1). La statistique de test serait $S = \hat{\sigma}^2 = \frac{1}{n-1} \sum (X_i - \bar{X})^2$. Le test aura la forme $T = \mathbb{1}_{\hat{\sigma}^2 < t}$.

Question 19. *Comment faire un plan de réponse à un exercice de test ?*

On reprend le même plan qu'auparavant en changeant le point 3.

1. Identifier le cadre
   -Identifier les paramètres connus et les paramètres inconnus. (Certaines fois, la forme de la loi elle-même est inconnue.)
   -On formule des **hypothèses** sur les paramètres (ou loi) inconnus.
2. Traduire l'énoncé en quantité(s) dépendant(es) des X_i
   -On construit les estimateurs naturels des paramètres qui apparaissent dans les hypothèses.
   -On recherche une quantité (la fameuse statistique de test S) qui a deux comportements différents que l'on soit sous H_0 ou sous H_1. (C'est ici qu'il faut réfléchir !)
3. On construit le test
   -On regarde comment se comporte S sous H_0. (Le $\bar{X}$ par exemple)
   -Le test aura très souvent la forme $T = \mathbb{1}_{S > t}$ (ou $T = \mathbb{1}_{S < t}$ ou $T = \mathbb{1}_{|S| > t}$, etc...).
   -A nous de trouver t tel que l'erreur sous H_0, qui est $\sup_{\theta \in H_0} \mathbb{P}_\theta(T = 1)$, ne dépasse pas la valeur seuil de l'énoncé (le α).
4. Conclure

Question 20. *Qu'est-ce qu'un test du χ^2 ?*

Ce test permet de vérifier que la collection de variables $X_1, \ldots, X_n$ suit une loi donnée. Attention ! Ce test ne fonctionne que lorsque la loi donnée est une **loi discrète de nombre d'états fini**. On note $1, \ldots, k$ les différents états. La loi testée est donc de la forme : $\mu_j = \mathbb{P}(X = j)$. Les hypothèses H_0 et H_1 auront toujours la même forme.

$$H_0 = \text{“Les } X_i \text{ suivent la loi } \mu\text{”} \quad H_1 = \text{“Les } X_i \text{ ne suivent pas la loi } \mu\text{”}$$

La démarche à suivre ensuite doit être apprise par coeur.